ARMY RESEARCH LABORATORY

Effects of Multimodal Displays About Threat Location on Target Acquisition and Attention to Visual and Auditory Communications

by Monica M. Glumm, Kathy L. Kehring, and Timothy L. White

ARL-TR-4074 **April 2007**

NOTICES

Disclaimers

Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5425

ARL-TR-4074 **April 2007**

Effects of Multimodal Displays About Threat Location on Target Acquisition and Attention to Visual and Auditory Communications

Monica M. Glumm, Kathy L. Kehring, and Timothy L. White
Human Research and Engineering Directorate, ARL

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
April 2007		

4. TITLE AND SUBTITLE	
Effects of Multimodal Displays About Threat Location on Target Acquisition and Attention to Visual and Auditory Communications	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 62716AH70
Monica M. Glumm, Kathy L. Kehring, and Timothy L. White (all of ARL)	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory Human Research and Engineering Directorate Aberdeen Proving Ground, MD 21005-5425	8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-4074
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This laboratory experiment examined the effects of paired sensory cues that indicate the location of targets on target acquisition performance, the recall of information presented in concurrent visual and auditory communications, and perceived workload. The multimodal cueing techniques assessed in this study were Visual+Spatial Language, Visual+3-D Audio, Visual+Tactile, and Spatial Language+Tactile. A unimodal "visual only" cue was included as a baseline. Except for reaction times to cues, no significant differences were found between the multimodal cue conditions and the Visual Only mode in primary and secondary task performance or subjective workload. Reaction times were faster in the Visual+3-D Audio and the Visual+Tactile conditions than in modes that included a spatial language cue. Reaction times to the visual+spatial language cue were faster than the spatial language+tactile cue, but no significant differences were found between the Visual+Spatial Language and the Visual Only modes. Adding the 3-D audio cue to the visual cue significantly improved reaction time beyond that of the Visual Only condition, but no significant difference was found between the Visual Only and the Visual+Tactile modes. Reaction times to cues were slower when communications were presented visually, but no interaction was found between communications modality and cue condition on this measure. Communications modality, however, did have a different effect on subjective ratings of effort in the Visual+Tactile mode than in the other cue conditions. In the Visual+Tactile mode, ratings of effort were significantly lower when communications were presented auditorily than when they were presented visually, but communications modality did not appear to affect ratings of effort in the other cue conditions.

15. SUBJECT TERMS

display; sensory modalities; target acquisition

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Monica M. Glumm
Unclassified	Unclassified	Unclassified	SAR	55	19b. TELEPHONE NUMBER *(Include area code)* 410-278-5986

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

INTENTIONALLY LEFT BLANK

Acknowledgments

The authors would like to express their appreciation to the following colleagues and friends at the U.S. Army Research Laboratory (ARL) who assisted in the conduct of this research:

SGM Enoch Godbolt for his assistance in obtaining the military volunteers for this study,

SSG Luis Rodriguez for transporting the Soldiers to and from the airport, motel, and ARL,

Nellie Moore for coordinating with the mess hall and ensuring that the Soldiers received a warm breakfast and lunch each day during their visit,

Tim Mermagen and Theo Feng for their assistance in the recording of the situation reports,

Dr. Paula Henry and 1LT Kara Delaney for providing hearing tests to the participants, and

Patty Burcham for her support in the conduct of the investigation.

Last, but never least, the authors would like to thank the Soldiers who participated in this research for their enthusiasm and dedication to the research tasks.

INTENTIONALLY LEFT BLANK

1. Introduction

Future combat vehicle systems will be lighter in weight and the size of their crews will be potentially reduced from three to two. Communication and target acquisition tasks that are currently shared by two crew members could become the responsibility of one: the commander-gunner. Task and workload analyses have shown that the commander-gunner will often become overloaded when he attends to auditory or digital communications while scanning for threats and engaging targets (Mitchell, Samms, Glumm, Krausman, Brelsford, & Garrett, 2004).

According to the multiple resource theory, two tasks can be time shared more efficiently if these tasks do not share the same perceptual modality and employ different working memory codes (Wickens, 2002). Although different sensory modalities may be used in the acquisition of targets and information presented in radio transmissions, both tasks can impose high demands on cognitive resources and attention. Target detection and identification require a search through long- and short-term memory stores, as do the interpretation and prioritization of auditory or visual information about the battlefield. If information about the battlefield is to be recalled, it must be rehearsed to be maintained in memory (Wickens & Hollands, 2000). As task demands increase and attentional resources employed in these processes are depleted, performance of one or both tasks is expected to decline.

As technologies advance, Soldiers will be provided far more information about the battlefield than they are provided now. The combat advantage that information can provide is well recognized, but as the amount of information the Soldier must process increases, so does the potential for cognitive overload. In an effort to reduce the burden on over-taxed resources and increase the speed and accuracy with which Soldiers capture, interpret, and act upon information, Army researchers are investigating different techniques for displaying information that Soldiers have identified as critical. High on Soldiers' lists of information requirements is enemy location.

Information about the location of enemy threats or targets might be provided in an auditory, visual, or tactile modality. Auditory cues might be presented verbally in spatial language (e.g., "5 o'clock") or in three-dimensional (3-D) audio sounds that appear to emanate from the location of the target in space. Research about collision avoidance in aircraft has found reductions in target acquisition time and perceived workload when target location cues were provided in 3-D audio (Begault, 1993; McKinley, Erickson, & D'Angelo, 1994; McKinley et al., 1995; Simpson et al., 2004). Improvements in performance and workload have also been found when 3-D audio and visual cues were paired (Tannen et al., 2000). In one such study, Simpson et al. (2004) measured times to visually acquire targets in a simulated flight task in four display conditions. One of the four conditions provided no information about the presence or location of targets (No Display). In a second condition, a non-spatialized auditory alert (i.e., "chirp") was provided to signal the presence of a target and was accompanied by a visual traffic advisory system (TAS) display that showed the

relative altitude and direction of the target (Visual+Audio Alert). In the third condition, the visual display was supplemented by a verbal cue that provided target location information in clock positions (Visual+Clock-Coordinate). A spatialized audio chirp indicating target direction and a verbal cue indicating relative altitude, supplemented the visual display in the fourth condition (Visual+3-D Audio). The researchers found that target acquisition times were slower in the first two conditions (i.e., No Display and Visual+Audio Alert), which were also the only conditions in which targets were undetected. However, pairing the auditory alert about the presence of a target with the visual display showing target location significantly reduced target acquisition times over the No Display condition. Simpson et al. (2004) also found an average 25% reduction in target acquisition time with the Visual+3-D Audio mode compared to the other conditions they studied. Target acquisition time was 1 second slower in the Visual+Clock Coordinate condition than in the Visual+3-D Audio mode. The researchers concluded that response times to a visual TAS could be significantly reduced if a non-spatialized clock-coordinate audio display were added; however, even greater reductions in response time could be achieved if the visual display were supplemented with a spatialized 3-D audio display.

Research has shown that tactile displays may offer a viable alternative to auditory and visual displays of information particularly in situations when the auditory and visual channels are heavily loaded. Tactile cues have been found useful in alerting operators to unexpected, high priority events (Calhoun, Draper, Ruff, Fontejon, & Guilfoos, 2003) and an acceptable substitute for auditory alerts as a redundant cue to visual alerts (Calhoun, Fontejon, Draper, Ruff, & Guilfoos, 2004; Krausman et al., 2007). Calhoun et al. (2004) found that in situations when visual load was high, reaction times to tactile cues were faster than reaction times to visual alerts and as fast as those to auditory cues. Like auditory cues, tactile cues do not require the operator to look in any particular direction to receive the information. Research suggests that tactile cues can enable the performance of multiple tasks better than visual cues. In one study, a tactile cue elicited faster reaction times in the detection of system faults than a visual cue and was less disruptive in the performance of concurrent tracking and visual monitoring tasks (Calhoun, Draper, Ruff, & Fontejon, 2002). In another study, the performance of helicopter pilots in counteracting the effects of drift was less affected by a secondary cognitive task when a tactile display was used to provide direction information (Van Erp, Veltman, Van Veen, & Oving, 2003). In other investigations, improvements in navigation performance and reductions in workload were found with tactile displays compared to visual displays (Chiasson, McGrath, & Rupert, 2003; Van Erp, Meppelink, & Van Veen, 2002).

The present study was the third in a series of investigations conducted in support of a Situational Understanding Army Technology Objective (SU ATO). The objective of this ATO is to “develop, demonstrate, and transition unit of action (UoA) Soldier information system interface guidelines that facilitate Soldiers gaining situational understanding and enable planning and acting within the adversary’s decision cycle.”

The first study examined the effects of one visual cue and two types of auditory cues (i.e., spatial language and 3-D audio) about threat location on target acquisition performance and the recall of information presented in concurrent auditory communications (Glumm, Kehring, & White, 2005). Two baseline conditions were also included in the analyses: no cues (Baseline 1) and target presence cues only (Baseline 2). The results of this study indicated that significantly less information was recalled from the auditory communications in Baseline 1 where participants were required to continuously scan for targets. Baseline 1 and 2 were both associated with slower target acquisition times. In modes in which target location cues were provided, 100% of the targets presented were acquired, compared to 94% in Baseline 1 and 95% in Baseline 2. No differences were found between the 3-D Audio and the Visual modes in target acquisition time or perceived workload. On average, target acquisition times were 1 second faster in the 3-D Audio and Visual modes than they were in the Spatial Language condition.

These latter findings were similar to those of Simpson et al. (2004) who did not present any concurrent verbal communications that could potentially interfere with the perception of the verbal cues about target location. Therefore, it was believed that the 1-second difference in target acquisition time between the Spatial Language condition and the Visual and 3-D Audio modes may have been attributable to other factors. First, the delay in target acquisition in the Spatial Language mode may have been influenced by the structure of the verbal cue. As in the study by Simpson et al. (2004), words that defined the location of the target were presented in the latter half of the cue, preceded by a verbal alert (e.g., “Target -- 9 o’clock” or “Traffic -- 9 o’clock high”). A 1-second delay in the receipt of information about the location of a target might be expected to result in a similar delay in the time to acquire the target. However, Glumm et al. (2005) also believed that increases in target acquisition time in the Spatial Language mode might be attributable to additional processing requirements. When Haas, Pillalamarri, Stachowiak, and Lattin (2005) presented target location information in plus and minus degrees (e.g., “Target -- minus 15 degrees”) no differences were found between the Verbal and the 3-D Audio modes. Glumm et al. (2005) surmised that the verbal cues employed by Haas et al. (2005) may have provided a more immediate indication of whether the target lay to the left (i.e., “minus”) or right (i.e., “plus”) of 0 degrees as might a 3-D audio sound that appears to emanate from one or the other direction in space.

The second study by Glumm, Kehring, and White (2006) extended the evaluation of cueing techniques to the tactile mode and included communications modality as a second factor. In this study, the participants monitored auditory or visual communications while performing target acquisition tasks in each of five cue conditions: (1) Baseline (2) Visual (3) Spatial Language (4) 3-D Audio and (5) Tactile. In order to minimize differences between cues in the time in which target location information was presented, the researchers restructured the spatial language cue by replacing the word “Target!” with information about the location of the target (i.e., “9 o’clock … 9 o’clock”). Measures of reaction time (i.e., time from cue presentation to movement of the crosshairs) and time to slew to target (i.e., time to first shot minus reaction time) were also obtained to gain further

insight into factors that might contribute to differences among cueing techniques in target acquisition time. In this second study, targets were well camouflaged and less conspicuous than they were in the initial investigation to encourage the participant to use the cue to localize the target to a specific clock position rather than merely slew in the direction indicated. Thus, the speed at which a target was acquired would depend primarily on the fidelity of the target location cue and the size of the area to be searched.

The results of this second investigation indicated a much stronger effect of target location cues on target acquisition performance, most likely because of the decrease in the visibility of the targets. Participants hit an average of 98% of the targets presented when cued about the location of targets, compared to 64% in the Baseline condition (no cues). When target location cues were provided, (a) time to first shot was an average 26% faster, (b) 30% more information was recalled from the auditory and visual communications, and (c) overall workload scores were 17% lower. The analysis did not reveal a significant difference between the Visual and Spatial Language conditions in time to first shot. However, time to first shot in these latter two modes was an average 13% faster than in the Tactile condition and 26% faster than in the 3-D audio mode. The results of the analyses of reaction time to the target location cues and time to slew to target revealed a significant effect of cue condition. Reaction time was slower in the Spatial Language mode than in the Visual, 3-D Audio, and Tactile conditions; however, time to slew to the target was an average 24% faster in the Visual and Spatial Language modes than in the 3-D Audio and Tactile conditions. Overall workload scores were 14% higher in the 3-D audio mode than in the other cue conditions. Communications modality did not have a significant effect on the amount of information recalled from the communications or on target acquisition performance. No interactions were found between communications modality and cue condition. However, on average, 10% more information was recalled from communications when target location cues were provided in the Visual mode than in the other cue conditions.

The reduction in the conspicuity of the targets in the second experiment appeared to have the greatest effect on target acquisition performance in the Baseline and 3-D Audio conditions. In the first study, participants hit 100% of the targets presented in the 3-D Audio mode compared to 93% in the second investigation. The percentage of hits achieved in the Baseline condition dropped from 94% to 64%, whereas 100% of the targets presented in the Visual and Spatial Language modes were hit in both studies. In the second investigation, mean time to first shot increased by more than 3 seconds in both the 3-D Audio mode and the Baseline condition. In the Visual mode, time to first shot increased by approximately 1 second; however, no difference was found between the two studies in time to first shot in the Spatial Language mode. It is believed that the restructuring of the spatial language cue eliminated the delay in the receipt of information about target location and helped to offset an increase in target acquisition time similar to that found in the Visual mode. The small but statistically significant difference that remained between the Spatial Language mode and the other cue conditions in reaction time may have reflected the time involved in "converting the utterance linguistically into meaning" (Loomis, Lippa, Klatzky, & Golledge, 2002).

The purpose of the present and third investigation in the series was to quantify the effects of combining cueing techniques on Soldier performance and workload. One unimodal (i.e., Visual Only) and four multimodal cues (i.e., Visual+Spatial Language, Visual+3-D Audio, Visual+Tactile, and Spatial Language+Tactile) were evaluated. The following were hypothesized for this investigation:

H1: Reaction time (time from cue presentation to movement of crosshairs) is expected to be faster in the Visual+Tactile and Visual+3-D Audio conditions because the tactile and the 3-D audio cues will arouse attention and both cues in each pair provide an immediate indication of the direction to slew to acquire the target. The spatial language cue does not provide an immediate indication of slewing direction, nor does it have the attentional value of the 3-D audio and tactile cues because of the potential for interference by the visual and auditory communications. Therefore, reaction time is expected to be slower in the Visual+Spatial Language and the Spatial Language+Tactile conditions, but no significant differences are expected between these two modes and the Visual Only condition.

H2: Cue condition is not expected to affect time to slew to the location of the target (i.e., time to first shot minus reaction time) because each of the five conditions includes a cue that directly indicates the clock position of the target (i.e., visual and spatial language), thus, reducing uncertainty that can slow slewing speed.

H3: Time to first shot and time to hit are expected to be faster in the Visual+Tactile and the Visual+3-D Audio conditions because of a faster reaction time (see H1). No differences in time to first shot or hit are expected among the remaining cue conditions: Visual+Spatial Language, Spatial Language+Tactile, and Visual Only.

H4: More information will be recalled from situation reports (SITREPs) in the Visual Only mode than in the other conditions because fewer sensory signals are provided that can distract from the capture and rehearsal of the information presented.

2. Objective

The objective of this laboratory study was to measure and compare the effects of paired sensory cues (i.e., auditory, visual, and tactile) that indicate the location of targets on target acquisition performance and the recall of information presented in concurrent visual and auditory communications.

3. Method

3.1 Participants

The participants were 20 male Soldiers who ranged in age from 22 to 38 years (mean *[M]* = 28.1 years; standard deviation *[SD]* = 4.8 years) with 1 to 16 years of time in service (*M* = 7.7 years; *SD* = 4.4 years) and a similar amount of time in their military occupational speciality (MOS). The participants were commanders and gunners of the Bradley fighting vehicle (BFV) or the M1 tank with an MOS of 19D or 19K, respectively. All but one Soldier had completed at least one tour in Iraq and had seen combat during their tour(s).

All participants passed tests of color vision and met visual acuity requirements of 20/20 in one eye and at least 20/30 in the other eye, corrected or uncorrected. The hearing threshold levels (HTL) of the participants corresponded to Army physical profile H2 which specifies an average HTL of no more than 30 dB, no individual HTL greater than 35 dB at 500, 1000, and 2000 Hz, and no HTL greater than 55 dB at 4000 Hz (U.S. Army, 1991). The participants had otoscopically normal ears (i.e., no blockage or infection) and no history of otologic pathology (i.e., hearing problems) as reported by the participants.

The voluntary, fully informed consent of the persons used in this research was obtained as required by 32 Code of Federal Regulations (CFR) 219 and Army Regulation (AR) 70-25 (1990). The investigators adhered to the policies for the protection of human subjects as prescribed in AR 70-25.

3.2 Apparatus

3.2.1 Control Station and Target Scenario

The participant's control station consisted of a computer monitor (17-inch[1], Dell Trinitron[2]) and a joystick manufactured by Saitek (Cyborg 3D Rumble Force Stick). The participant sat approximately 25 inches from the computer monitor (i.e., seat reference point to screen). The monitor presented a 10-degree horizontal view[1] of the 360-degree field about the vehicle in which the participant was theoretically operating. The joystick controlled the movement of the scene behind crosshairs, which were fixed in the center of the visual display. The participant was able to scan the 360-degree field around the vehicle by twisting the joystick to the left or the

---

[1]This was based on discussions with United Defense Limited Partnership (UDLP) who is responsible for the design of the crew station in the infantry carrier vehicle (ICV). UDLP provided information about the field of view of the commander's independent viewer (daylight TV sensor) in the BFV A3 (i.e., wide field of view [WFOV]: 10 degrees x 7.5 degrees) and best guesses about the FOV of the sight in the ICV (9 degrees circular). At the time of this study, there had not been any decision regarding the FOV of the commander-gunner's sight in the ICV or the size of the flat panel upon which the sight image would be displayed. A best guess was that the sight image would be presented on one of the main 15-inch flat panel displays and, if desired, on a smaller square flat panel called the crewman's remote interface system (CRIS).

[2]Trinitron is a registered trademark of Sony.

right. The farther the hand control was twisted, the faster the movement of the target scene. Movement of the scene behind the crosshairs was limited to the horizontal plane. An azimuth indicator was provided at the bottom of the target scene to show the direction in which the main gun was pointing (see figure 1). The indicator was similar to that incorporated along the turret ring of the BFV; however, the azimuths were provided in clock positions rather than in mils to preclude the need for conversions.

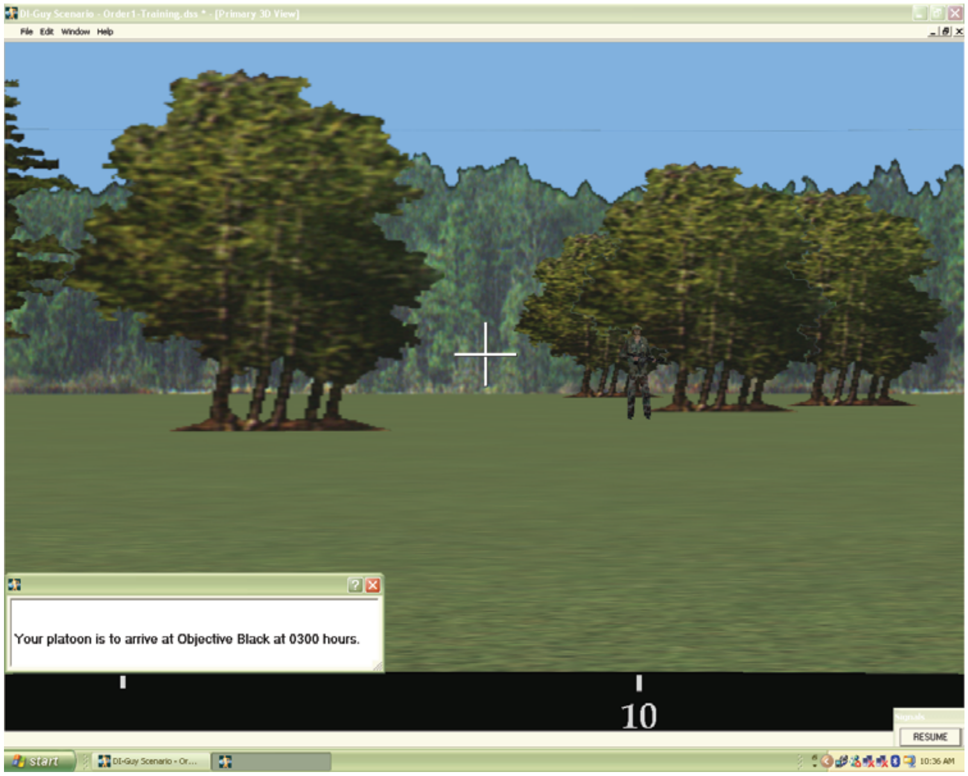


Figure 1. Target scene with azimuth indicator and SITREP segment.

Each target was an individual dismounted Soldier presumed to be an enemy. This choice of target type was based on a prioritized list of critical information requirements and related threats identified by subject matter experts. All targets were located at a distance of 75 meters from the participant's vehicle. The targets were equal in size and presented along the vertical centerline of the visual display.

A hit on the target was recorded when the trigger on the joystick was pulled while the crosshairs were on any portion of the target. The target fell to the ground to indicate to the participant that a hit had been scored. The DiGuy Scenario[3] (Version 5.2.3) was used in the development of the target scenarios, the presentation of the target cues, the interpretation of the input from the joystick, and data collection.

[3]DiGuy Scenario is a trademark of Boston Dynamics.

3.2.2 Target Location Cues

One unimodal and four multimodal cues about target location were examined in this study. These five cue conditions were (1) Visual Only, (2) Visual+Spatial Language, (3) Visual+3-D Audio, (4) Visual+ Tactile, And (5) Spatial Language+Tactile. The two cues in each of the four pairings were presented simultaneously. All cues were 2.5 seconds in duration and their presentation was controlled by computer. Target location cues were presented once for each target presented, relative to the 12-o'clock position. The following describes these cues and associated apparati:

(1) Visual Only. The visual cues about target location were provided by an icon that resembled a one-handed clock without numbers (see figure 2). The direction in which the hand on the clock was pointing indicated the location of the target within the 360-degree field about the vehicle platform. Target locations were incremented in hours as were the spatial language, 3-D audio, and tactile cues. Ten clock positions were used. No targets were presented at the 12-o'clock or the 6-o'clock positions, partly to avoid front-back reversals that can occur when these cues are presented in the 3-D audio mode (Begault, 1991). The visual icon was 2.5 by 2.5 inches in size and appeared at the bottom center of the screen just above the azimuth indicator.

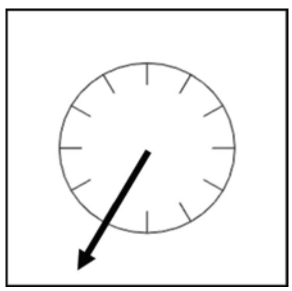


Figure 2. Icon providing directional cue in the visual modality.

(2) Visual+Spatial Language (Speech). The visual cue just described was paired with a verbal cue that was presented in a clock-type format. An example of this type of spatial language cue is "5 o'clock…5 o'clock." Like the visual cue, the total duration of the verbal cue was 2.5 seconds. The spatial language cues were pre-recorded in a female voice in contrast to the auditory communications that were presented in a male voice.

(3) Visual+3-D Audio (Non-Speech). The visual cue was paired with a spatialized, non-speech audio signal. Each audio signal was 2.5 seconds in duration and consisted of two, 1.0-second tones that were 0.5 second apart. The tones were spatialized when they were played through a Veridian Engineering 3-DVALS[4] audio sound engine so that each tone appeared to emanate from a different clock position. The spatialization used generic head-related transfer functions (HRTF) developed by the U.S. Air Force using a Kelso Electronic Mannequin for Auditory Research acoustic head.

[4]3-DVALS is a registered trademark of Veridian Corporation.

(4) Visual+Tactile. The visual cue was paired with a tactile signal transmitted by one of ten electromechanical vibrators or tactors. The tactor system was developed by the Massachusetts Institute of Technology (MIT) under the Collaborative Technology Alliance (Lockyer, 2004). Each tactor was approximately 1.5 cm in diameter and was connected by wire to a remote unit that incorporated a 9-volt battery. The tactors were incorporated into a belt that the participant wore around his torso. The belt was adjusted with Velcro[5] to accommodate the circumference of the participant's torso. The belt had 24 small pockets in which the tactors were placed to align them with each of the ten clock positions. The belt was attached to suspenders that could be adjusted to maintain the belt at the desired position. Each tactile signal consisted of two, 1.0-second vibrations presented 0.5 second apart for a total duration of 2.5 seconds. The vibrations were low in intensity and similar to those emitted by a cellular phone.

(5) Spatial Language+Tactile. This cue paired the verbal cue presented in a clock-type format and the tactile signal described before.

3.2.3 Communications and Questionnaires

The auditory and visual communications presented during the target engagement periods were in the form of a SITREP. The auditory communications were pre-recorded in a male voice and presented monaurally to the participant through stereo earphones (Sony MDR[6] 7506). All auditory cues and communications were normalized to 21 dB, and noise reduction was applied. The intensity levels of the auditory cues and communications were the same as those in the preceding investigation (Glumm et al., 2005). These intensity levels, as measured through an artificial ear, were 73 peak dB for the spatial language cue and 78 peak dB for the 3-D audio cue. The decibel peaks for the auditory communications ranged from 65 to 70. Given that all auditory cues were easily detected above the auditory communications, audiologists do not consider the differences in decibel peaks between the spatial language and the 3-D audio cues to be of any consequence.

The visual communications were displayed in a text box at the lower left of the target scene in Microsoft Sans Serif 12-point bold font (see figure 1). Each SITREP consisted of 15 individual sentences or segments (see appendix A). The number of words in each sentence ranged from 6 to 13. Each sentence was presented one at a time for 5 seconds with a 5-second time interval between sentence presentations. The time to speak the longest 13-word sentence at a normal rate of speech ranged from approximately 4 to 5 seconds. Given a maximum sentence length of 13 words, a display time of 5 seconds accommodated a reading speed of approximately 156 words per minute (2.6 words per second) and was below the average 8th grade reading speed of 200 words per minute (3.3 words per second) (Boff & Lincoln, 1988). The 15 sentences in each SITREP contained 20 items of information about the current battlefield. Ten of the sentences contained one item of information, and five of the sentences contained two items. Ten of the 20 items of information (50%) changed in each SITREP presented. The total duration of each

---

[5]Velcro is a registered trademark of Velcro USA, Inc.
[6]Not an acronym

SITREP was 2.5 minutes (150 seconds). The first sentence in the SITREP was presented 5 seconds after the start of each target set. Twenty-five SITREPs were prepared and pre-recorded: five SITREPs for training in each of the five cue conditions and 20 SITREPs for testing (i.e., two target sets in each of the ten experimental conditions).

After the completion of each target set, the participant was asked to complete a questionnaire that consisted of ten questions pertaining to the information contained in the SITREP. An example of this questionnaire is provided in appendix B. The answers to each question were written and required a one- or two-word written response. Each answer was worth 2 points. If the participant did not provide an answer to a question or the answer was wrong, the participant scored 0 points. If the participant omitted a word from an answer that required two words or if one of the words in the answer was wrong, the participant scored 1 point. The participant had a maximum of 3 minutes to answer the ten questions. The order of presentation of the SITREPs and associated questionnaires was counterbalanced.

3.2.4 The National Aeronautics and Space Administration (NASA) Task Load Index (TLX)

The NASA-TLX was used to assess the participant's experience of workload (Hart & Staveland, 1988). This technique uses rating scales to assess mental, physical, and temporal demands, performance, effort, and frustration. Initially, each of these six workload factors is assigned a weight based on the responses of the participant to pairwise comparisons. In these comparisons, the six factors are presented in 15 possible pairs, and for each pair, the participant is asked to circle the factor that s/he perceives contributed most to his or her workload experience. The participant then completes rating scales that provide a measure of the magnitude of the workload for each factor. Those factors perceived by the participant to have contributed most to his or her workload experience are given more weight in computing an overall workload score. The paired comparisons worksheets and the workload rating scale are provided in appendix C.

4. Procedures

4.1 Experimental Design

The study was a 5x2, fixed factor design with cue condition and communications modality all as within-subject effects. The five cue conditions that were evaluated in this study were (1) Visual Only, (2) Visual+Spatial Language, (3) Visual+3-D Audio, (4) Visual+Tactile, and (5) Spatial Language+Tactile. In all cue conditions, the participant's crosshairs automatically returned to the 12-o'clock position after each target presentation. The participant was not able to move his crosshairs from that position until cued about the location of another target.

The primary task of the participants was to find and engage targets as quickly as possible. Their secondary task was to recall information contained in SITREPs that were presented in one of two

communications modalities throughout each target set. The two SITREP communications modalities were (1) auditory and (2) visual.

The dependent variables in this study included measures of primary and secondary task performance and subjective ratings of workload. In the target acquisition task, the dependent variables were time to first shot and the percentage of hits. For those targets hit, time to hit was also recorded. The time to first shot and the time to hit were calculated from the time at which the target was presented to the time of trigger pull. Other measures of target acquisition performance included reaction time (i.e., time from cue presentation to movement of the crosshairs) and time to slew (i.e., time to first shot minus reaction time). The dependent variable in the secondary task was the total number of points scored on the SITREP questionnaires administered after each target set in each cue condition. Overall workload scores were computed from subjective ratings on each of six dimensions of workload obtained with the NASA-TLX.

One participant was trained and tested at a time. The duration of training and testing for each participant was approximately 5.7 hours. The procedures followed for each participant are described next.

4.2 Training

Each volunteer was briefed about the purpose of the investigation, the procedures to be followed during the study, and any risks involved in his participation. The investigator read the volunteer agreement affidavit aloud to the participant, who followed along. If the participant agreed to participate in the investigation, he completed the information on the last page of the affidavit and signed it. A demographic questionnaire was then administered to the participant to obtain pertinent information about his background (see appendix D).

A vision tester manufactured by Titmus Optical Company, Inc., was used to assess the participant's vision at near and far distances to ensure that the participant met visual acuity requirements of 20/20 in one eye and at least 20/30 in the other eye, corrected or uncorrected. The participant was also required to pass a test for color vision. A certified audiologist administered a hearing test with an AC40 clinical audiometer manufactured by Interacoustics A/S. The participants were required to have HTLs corresponding to Army physical profile H2 (U.S. Army, 1991), otoscopically normal ears, and no history of otologic pathology.

Each participant was trained and tested in performing the target acquisition tasks in each of the five cue conditions in one of the two communications modalities before being trained and tested in the other. The order of training and testing in each cue condition and communications modality was counterbalanced.

During the initial training period, the participants were instructed in rating their workload experience using the NASA-TLX. They also received practice in localizing the tactile and 3-D audio cues presented at the ten clock positions. For each of these modes, the investigator presented cues at each of the ten clock positions, starting at the 1-o'clock position and ending at 11 o'clock.

At each cue presentation, the clock position of the cue was displayed on the participant's monitor. This process was repeated two more times. The investigator then presented the cues at each clock position in a randomized order to the participant who identified the clock position of the target. This latter process was repeated two more times.

During training in each communications modality, the participant was reminded that his primary task was to find and engage all targets as quickly as possible. If the target did not fall at trigger pull, the target had not been hit and the participant was required to re-engage. The participant was told that his secondary task was to attend to information contained in the SITREP and that he would be asked to recall information from the SITREP after each target set. He was informed that some of the details in the SITREP would change in each SITREP presented and that he should not rely on his memory of information contained in previous SITREPs.

Training in each communications modality included the completion of one run in each of the five cue conditions for a total of ten runs, and practice in using the NASA-TLX in rating the workload experience. Each run consisted of two target sets. Each target set was followed by the questionnaire that assessed the participant's knowledge of the information contained in the SITREP.

4.3 Testing

After a 15-minute rest break, the participant completed one test run in each of the five cue conditions in the communications modality in which he was just trained. Before each run, the participant was reminded that his primary task was to find and engage each target as quickly as possible and that his secondary task was to remember information contained in the SITREP.

As during training, each test run consisted of two target sets. Each target set consisted of five targets for a total of ten target presentations. The targets were presented once in each of the ten clock positions in a random order. Ten different random orders had been developed for testing and the presentation of these orders was counterbalanced among the experimental conditions. Each target was presented for a maximum duration of 20 seconds. Regardless of whether the target was hit, the target disappeared from the screen after the 20 seconds elapsed. The time at which the first target was presented at the beginning of a target set and the time between subsequent target presentations was varied to reduce expectancy. The time intervals between target presentations were randomized in 4-second increments ranging from 4 to 20 seconds, and these orders were counterbalanced among the experimental conditions. The start of a time interval between target presentations began 20 seconds after the preceding target had been presented, regardless of the time in which the participant scored a hit on that target. The total duration of each target set was 2.6 minutes (160 seconds).

Each target set was followed by a questionnaire assessing the participant's knowledge of the information contained in the SITREP that he had just received. After each run in each cue condition, the participant was asked to rate his workload experience using the NASA-TLX.

A 1-hour lunch break was provided before completion of training and testing in the second communications modality. At the conclusion of testing in all conditions, the participant completed a questionnaire to obtain his opinions and preferences with regard to the conditions evaluated (see appendix E).

5. Results

5.1 Target Acquisition

5.1.1 Reaction Time

It was hypothesized that reaction time (i.e., time from cue presentation to movement of crosshairs) would be faster in the Visual+Tactile and Visual+3-D Audio conditions because the tactile and the 3-D audio cues were less likely to conflict with the visual and the verbal communications and more likely to capture attention. It was also believed that these cues would provide a more intuitive indication of the direction to slew to acquire the target in the shortest distance. Reaction time was expected to be slower in the Visual+Spatial Language and the Spatial Language+Tactile conditions because some additional processing would be required to convert the verbal cue linguistically into meaning and into a mental image of the target's location. Reaction time in the Visual Only condition, however, was not expected to be different from reaction times in the latter two modes.

To test this hypothesis, a linear mixed model analysis was performed on reaction times in each of the five cue conditions for each of the two modalities in which the SITREPs were presented. The analyses revealed a significant effect of both cue ($F(4, 24) = 11.713$, $p < .001$) and communications modality ($F(1, 24) = 4.646$, $p = .03$), but no interaction was found between the two factors ($F(4, 24) = .210$, $p = .93$). The means and standard deviations in reaction time for each cue condition are shown in figure 3. Figure 3 also indicates the results of the *post hoc* analyses via the least significant difference (LSD) method, which are denoted by the letters shown within each bar in the chart. Like letters indicate that there are no significant differences between the cue conditions. The mean differences between cue conditions are provided in table 1.

As hypothesized, reaction times in both the Visual+3-D Audio and Visual+Tactile modes were faster than in the Visual+Spatial Language and the Spatial Language+Tactile conditions. Reaction times in the Spatial Language+Tactile mode were significantly slower than all other cue conditions, but reaction times in the Visual+Spatial Language mode were not significantly different from those in the Visual Only condition. The analysis did not reveal a significant difference between the Visual+3-D Audio and the Visual+Tactile modes, but reaction times were faster in the Visual+3-D Audio mode than in the Visual Only condition, and no significant difference was found between the Visual Only condition and the Visual+Tactile mode.

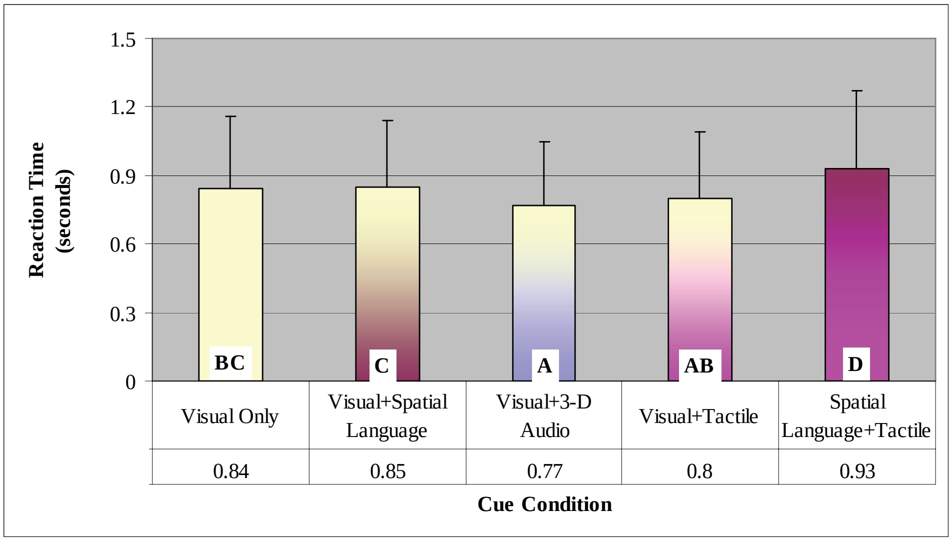


Figure 3. Mean reaction time by cue condition.

Table 1. Mean difference between cue conditions in reaction time.

Mode	Visual+Spatial Language	Visual+3-D Audio	Visual+Tactile	Spatial Language+Tactile
Visual Only	-.013 (*p*=.60)	**.072 (*p*=.005)***	.037 (*p*=.14)	**-.090 (*p*<.001)***
Visual+Spatial Language		**.085 (*p*=.001)***	**.051 (*p*=.05)***	**-.077 (*p*=.003)***
Visual+3-D Audio			-.035 (*p*=.17)	**-.162 (*p*<.001)***
Visual+Tactile				**-.127 (*p*<.001)***

Bold blocks indicate significant differences.

As shown in figure 4, reaction times were significantly slower when the SITREP communications were presented visually than when they were presented auditorily.

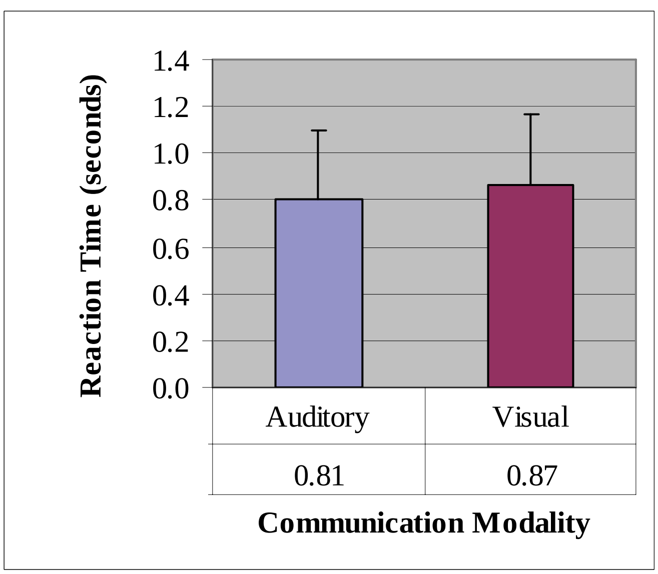


Figure 4. Mean reaction time by communications modality.

5.1.2 Time to Slew

Time to slew to the location of the target (i.e., time to first shot minus reaction time) was not expected to be affected by cue condition because each of the five conditions included a cue that directly indicated the clock position of the target (i.e., visual and spatial language), thus reducing uncertainty that could slow slewing speed. To test this hypothesis, a linear mixed model analysis was performed on time to slew. The results of the analysis did not reveal a significant effect of either cue ($F(4, 24) = 1.435$, $p = .22$) or communications modality ($F(1, 24) = .695$, $p = .41$), and no interaction was found between the two factors ($F(4, 24) = .410$, $p = .80$). The means and standard deviations in time to slew for each cue condition and communications modality are shown in figure 5.

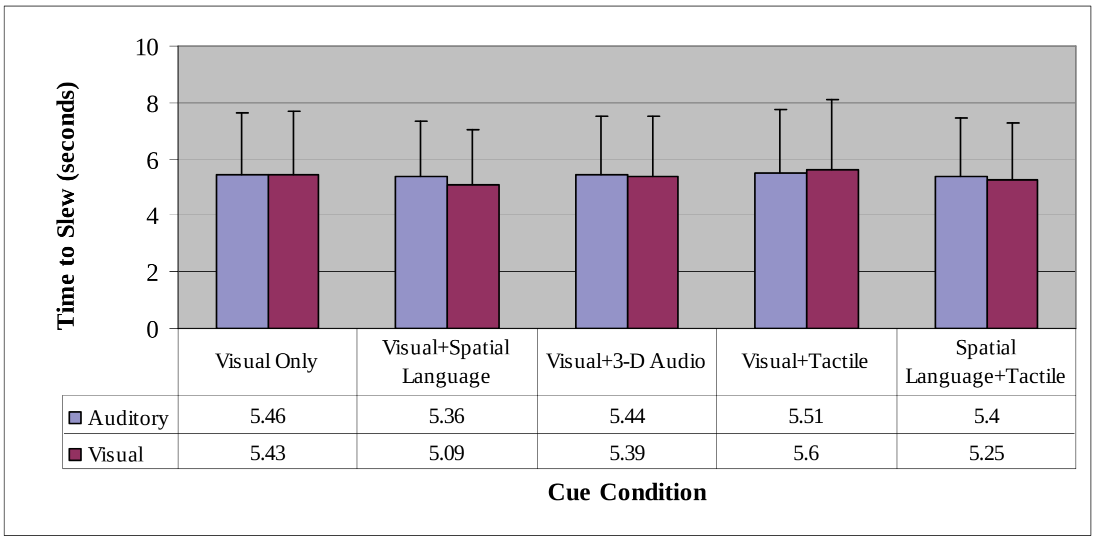


Figure 5. Mean time to slew by cue condition.

5.1.3 Time to First Shot

It was hypothesized that time to first shot would be faster in the Visual+Tactile and the Visual+ 3-D Audio modes than in the other cue conditions because of faster reaction times. No differences in time to first shot were expected between the Visual+Spatial Language, Spatial Language+Tactile, or the Visual Only conditions. A linear mixed model analysis was performed to test these hypotheses. Contrary to expectations, the analysis did not reveal a significant effect of cue ($F(4, 24) = 1.010$, $p = .40$) or communications modality ($F(1, 24) = .215$, $p = .64$), and no interactions were found between the two factors ($F(4, 24) = .459$, $p = .77$). The means and standard deviations for each cue condition and communications modality are shown in figure 6.

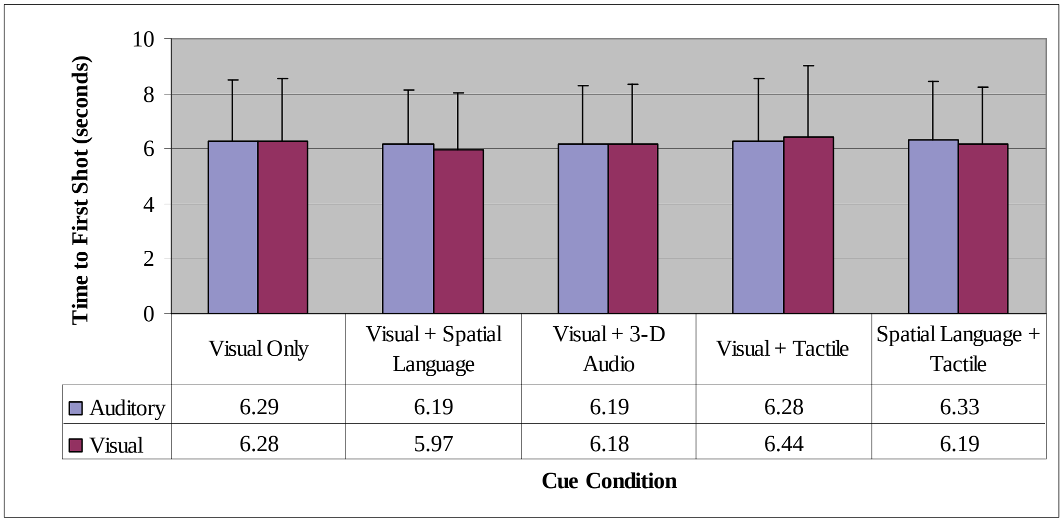


Figure 6. Mean time to first shot by cue condition.

5.1.4 Percent Hits

A linear mixed model analysis was performed on the percentage of hits achieved in each of the five cue conditions for each modality in which the SITREPs were presented. The results of the analysis failed to show a difference between cue conditions at the .05 level of significance (F(4, 24) = 2.268, p = .06). Nor did the analysis reveal a significant main effect of communications modality (F(1, 24) = 1.002, p = .32) or an interaction between cue and communications modes (F(4, 24) = 1.174, p = .33).

The means and standard deviations of the percentage of hits achieved in each cue condition by communications modality are shown in figure 7. Given that the significance level of differences between cue conditions is considered marginal, the results of *post hoc* analyses are of interest and are therefore included. The mean differences between cue conditions are provided in table 4.

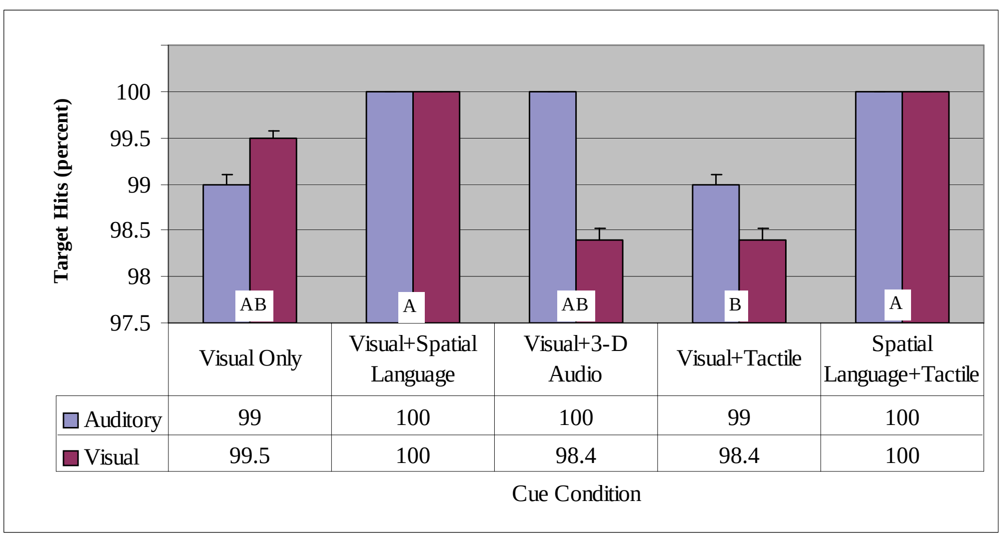


Figure 7. Mean percentage of target hits by cue condition.

Table 2. Mean differences between cue conditions in the percentage of targets hit.

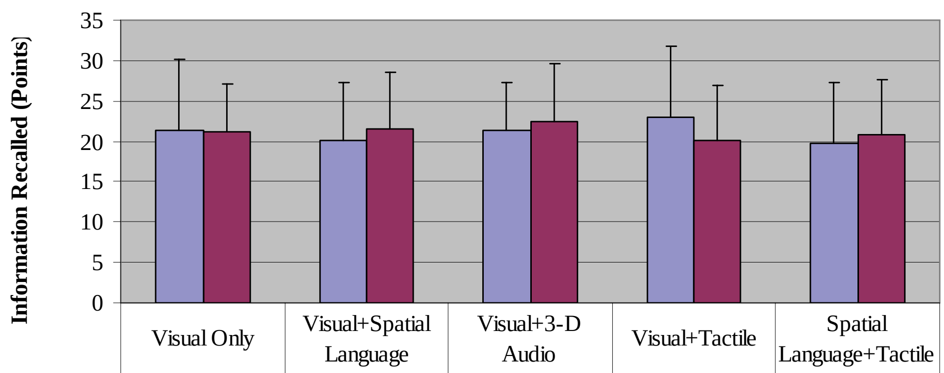
Mode	Visual+Spatial Language	Visual+3-D Audio	Visual+Tactile	Spatial Language+Tactile
Visual Only	-.008 (*p*=.16)	.001 (*p*=.92)	.005 (*p*=.32)	-.008 (*p*=.15)
Visual+Spatial Language		.008 (*p*=.13)	**.013 (*p*=.02)***	.000 (*p*=.98)
Visual+3-D Audio			.005 (*p*=.37)	-.008 (*p*=.12)
Visual+Tactile				**-.013 (*p*=.02)***

Bold blocks indicate significant differences.

The results of the *post hoc* analyses indicated that the percentage of hits achieved in the Visual+ Tactile condition was significantly lower than that in both the Visual+Spatial Language and the Spatial Language+Tactile modes.

5.2 Secondary Task Performance (scores on SITREP questionnaire)

It had been hypothesized that more information would be recalled from SITREPs in the Visual Only mode than in the other cue conditions because fewer sensory signals would be provided that would distract from the capture and rehearsal of the information presented. To test this hypothesis, a linear mixed model analysis was performed on the number of points scored on the SITREP questionnaires. The means and standard deviations for each cue condition by communications modality are shown in figure 8. Contrary to expectations, the results of the analysis did not reveal a significant effect of either cue ($F(4, 24) = .686$, $p = .60$) or communications modality ($F(1, 24) = .031$, $p = .86$), and no interaction was found between the two factors ($F(4, 24) = 1.430$, $p = .23$).



	Visual Only	Visual+Spatial Language	Visual+3-D Audio	Visual+Tactile	Spatial Language+Tactile
Auditory	21.4	20.0	21.3	23.0	19.8
Visual	21.1	21.6	22.5	20.1	20.8

Figure 8. Mean scores on SITREP questionnaire.

5.3 Subjective Workload

Separate linear mixed model analyses were performed on weighted ratings of workload on each of the six subscales of the NASA-TLX (i.e., mental, physical, and temporal demand, performance,

effort, and frustration). A linear mixed model analysis was also performed on overall weighted workload scores that were derived from ratings on each of the six subscales.

The mixed model analyses did not reveal a significant effect of cue on mental demand ($F(4, 24) = .867$, $p = .49$), physical demand ($F(4, 24) = .295$, $p = .88$), temporal demand ($F(4, 24) = .128$, $p = .97$), performance ($F(4, 24) = 1.643$, $p = .17$), effort ($F(4, 24) = 1.504$, $p = .20$), or frustration ($F(4, 24) = 2.319$, $p = .06$). No significant differences were found between cue conditions on overall weighted workload scores ($F(4, 24) = .325$, $p = .85$). No main effects of communications modality were found on mental demand ($F(1, 24) = 2.934$, $p = .09$), physical demand ($F(1, 24) = 2.377$, $p = .13$), temporal demand ($F(1, 24) = 1.060$, $p = .31$), performance ($F(1, 24) = .068$, $p = .80$), effort ($F(1, 24) = 1.098$, $p = .30$), or frustration ($F(1, 24) = .852$, $p = .36$). Here, too, no significant differences were found between modalities on overall workload scores ($F(1, 24) = 1.411$, $p = .31$). The analysis, however, did indicate an interaction between cue condition and communications modality on ratings of effort ($F(4, 24) = 4.577$, $p = .002$).

The interaction found between cue condition and communications modality for ratings of effort suggests that communications modality had different effects on the participants' perceptions of the amount of effort they expended in the Visual+Tactile mode than in the other cue conditions (see figure 9). In the Visual+Tactile mode, the participants perceived that they expended significantly less effort when the SITREPs were presented auditorily than when the SITREPs were presented visually. By comparison, the modality in which the SITREPs were presented did not have a significant effect on the participants' ratings of effort in the other cue conditions.

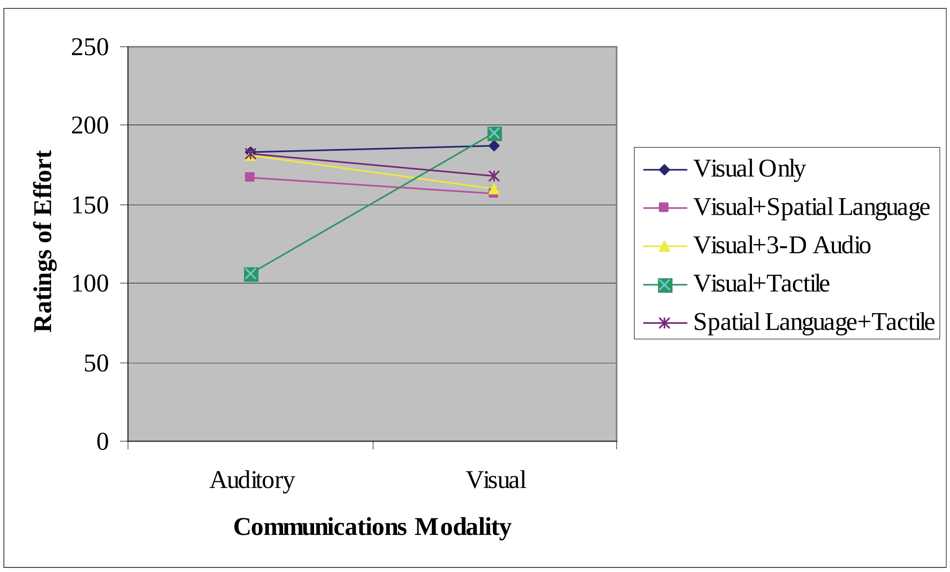


Figure 9. Interaction between cue condition and communications modality on subjective ratings of effort.

Given that the level of significance of differences between cue conditions in ratings of frustration is considered marginal, the results of *post hoc* analyses are of interest and are therefore shown in

figure 10 along with mean ratings and standard deviations. The mean differences between cue conditions are shown in table 3. *Post hoc* analyses indicated that participants experienced higher levels of frustration in the Spatial Language+Tactile mode than in the Visual+3-D Audio and the Visual+Tactile conditions. Ratings of frustration were also higher in the Visual Only condition than in the Visual+3-D Audio mode, but no other significant differences were found between cue conditions.

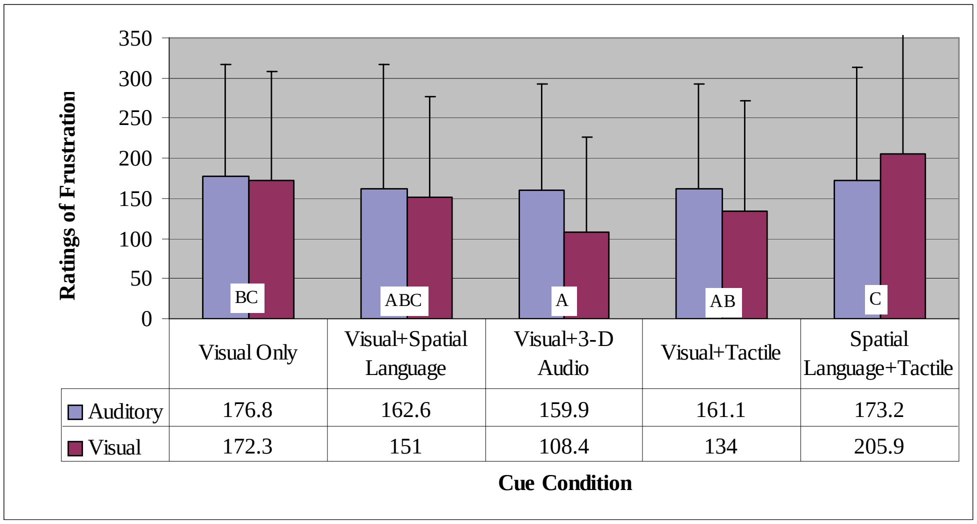


Figure 10. Mean ratings of frustration by cue condition.

Table 3. Mean differences between cue conditions in subjective ratings of frustration.

Mode	Visual+Spatial Language	Visual+3-D Audio	Visual+Tactile	Spatial Language+Tactile
Visual Only	18.865 (*p*=.36)	**40.400 (*p*=.05)***	27.000 (*p*=.19)	-15.025 (*p*=.46)
Visual+Spatial Language		21.535 (*p*=.30)	8.135 (*p*=.69)	-33.890 (*p*=.10)
Visual+3-D Audio			-13.400 (*p*=.51)	**-55.425 (*p*=.007)***
Visual+Tactile				**-42.025 (*p*=.04)***

Bold blocks indicate significant differences.

5.4 Participants' Preferences and Comments

Most participants ranked the Visual+3-D Audio cue condition as last with respect to how helpful the cue was in locating targets. Although the participants believed that the 3-D Audio cue was useful in providing directional information, they claimed that they had difficulty in using it to determine the exact clock position of the target. The Visual Only condition was ranked as third and the Spatial Language+Tactile condition was ranked as second. There was no difference in the ranking of the Visual+Spatial Language and the Visual+Tactile modes, both of which participants considered to be most helpful in locating targets.

Of those participants who claimed to rely more on one cue than another within a given cue pair, more participants indicated that they relied primarily on the spatial language cue in the Visual+

Spatial Language (92%) and the Spatial Language+Tactile conditions (85%). Some noted that they relied more on the spatial language cue in these latter modes because they were more accustomed to receiving target location information in that format. One participant noted that he did not use the tactile cue to locate targets in the Spatial Language+Tactile mode because he "did not want to take the time". The participants agreed that the 3-D audio and tactile cues were helpful in determining whether to slew to the left or to the right to acquire the target in the shortest distance, but only 15% of the participants claimed that they relied more on the 3-D audio cue in the Visual+3-D Audio mode. Many participants noted that they had difficulty localizing targets using the 3-D audio cue. In the Visual+Tactile mode, the percentage of participants who claimed to rely on one cue more than another was nearly split between the two cues with 54% claiming to have relied on the visual cue and 46% claiming to have relied on the tactile cue.

Although no significant differences were found between communications modalities in the performance of the secondary communications task, 75% of the participants believed that they were able to recall more information when the SITREPs were presented auditorily than when the SITREPs were presented visually. Only 25% of the participants preferred the visual communications modality to the auditory modality. Approximately 50% of the participants believed that the visual communications interfered with cue conditions that included a visual cue, and a similar percentage believed that the auditory communications interfered with cue conditions that included an auditory cue.

6. Discussion

The purpose of this investigation was to measure and compare the effects of paired sensory cues (i.e., auditory, visual, and tactile) indicating the location of targets on target acquisition performance and the recall of information presented in concurrent visual and auditory communications. A unimodal, Visual Only condition was included as a baseline to determine the extent to which a supplementary cue might influence changes in performance and workload.

6.1 Target Acquisition (Primary Task)

6.1.1 Reaction Time

It was hypothesized that reaction times would be faster in the Visual+3-D Audio and Visual+Tactile conditions because the 3-D audio and the tactile cues would provide a more intuitive indication of the direction to slew to acquire the target in the shortest distance. It was also hypothesized, based on previous research (i.e., Glumm et al., 2006), that reaction times would be slower in conditions that included a spatial language cue (i.e., Visual+Spatial Language and the Spatial Language+Tactile), but no differences were expected between these latter modes and the Visual Only condition.

In accordance with these hypotheses, the results of the analyses indicated that reaction times were faster in the Visual+3-D Audio and the Visual+Tactile conditions than in modes that included a spatial language cue. This finding may support the notion that the target location cues presented verbally in clock positions require some additional processing, which involves converting the words linguistically into meaning (Loomis et al., 2002). Contrary to expectations, however, reaction times to the visual+spatial language cue were faster than the spatial language+tactile cue, and no significant differences were found between the Visual+Spatial Language mode and the Visual Only condition. This may indicate that some participants may have relied more on the visual cue in the Visual+Spatial Language condition than they had perceived. Adding the 3-D audio cue to the visual cue significantly improved reaction time beyond that of the unimodal Visual Only condition, but no significant difference was found between the Visual Only mode and the Visual+Tactile condition. Although reaction times in the Visual+3-D Audio mode were not significantly faster than in the Visual+Tactile condition, it appears that the 3-D audio cue may have been more successful than the tactile cue in capturing attention.

Calhoun et al. (2004) found that reaction times to visual+aural and visual+tactile alerts were both significantly faster than reaction times to unimodal visual cues in a high noise environment. However, contrary to their expectations, they did not find a significant difference in reaction time between the two multimodal alert conditions. The researchers speculated that the "noise-robust visual cue" played a key role in reaction time, regardless of the modality of the cue with which it was paired, referencing claims by Colavita (1974) that the visual channel is normally sampled first and is allocated more attention.

Although about half the participants believed that the visual SITREP communications interfered with the visual cues and a similar number believed that the auditory communications interfered with the auditory cues, no interaction was found between cue condition and communications modality for any measure of target acquisition performance. However, the analysis did reveal a significant effect of communications modality on reaction time, which indicated that the participants responded more slowly to the target location cues when the SITREP communications were presented visually. This finding may indicate a potential conflict between the visual SITREPs and the visual search task where participants may have delayed movement of the crosshairs until they had finished reading the text message. Given the tendency for more attention to be allocated to the visual cue in the three multimodal conditions in which it was presented, slower reaction times to cues during visual communications may also reflect the time to shift attention spatially from the visual SITREPs to the visual cues (Wickens et al., 2000).

6.1.2 Time to Slew

For the present study, it was hypothesized that no differences would be found between cue conditions in time to slew because each condition included a cue that directly indicated the clock position of the target (i.e., visual or spatial language). As expected, the analysis did not reveal a significant difference between cue conditions on this measure, thus supporting the belief that

participants tended to use the visual or the spatial language cue to determine the clock position of targets rather than the 3-D audio or the tactile cue.

6.1.3 Time to First Shot

It was hypothesized that faster reaction times in the Visual+3-D Audio and the Visual+Tactile modes would result in faster times to first shot. Although reaction times were faster in the Visual+3-D Audio and the Visual+Tactile modes, times to first shot were not any faster in these than in the other cue conditions. Although no differences were found between conditions in time to slew, slewing speed in both the Visual+Spatial Language and the Spatial Language+Tactile mode may have been fast enough to compensate for the small but statistically significant increase in reaction times that were found in these latter modes. Note, however, that targets that were not acquired were not included in calculations of the mean time to slew to target or the time to first shot.

6.1.4 Percent Hits

Pairing the visual cue with the 3-D audio, tactile, and spatial language cues did not significantly increase the percentage of hits beyond that achieved in the Visual Only condition. However, the analyses indicated that the percentage of hits in the Visual+Tactile mode tended to be lower than in both the Visual+Spatial Language and the Spatial Language+Tactile conditions. This finding may reflect, in part, differences in the extent to which participants may have relied on one cue more than another within each pair and the fidelity of that cue in denoting the clock position of the target. More participants claimed to rely on the tactile cue in the Visual+Tactile mode (46%) than on the 3-D audio cue in the Visual+3-D Audio condition (15%). However, the tactile cue did not specify the location of the target as clearly as the spatial language cue and was less likely to be accurately recalled if the visual cue was missed.

Although neither the tactile nor the 3-D audio cue provided a direct indication of the clock position of the target, there was some evidence after localization training, which suggests that participants were able to locate targets better with the tactile cue than with the 3-D audio cue. Table 4 shows the average percentage of times the 20 participants correctly identified the clock position of 30 targets and the average error in localization using the two different cues after the initial training period.

Table 4. Target localization performance after training with 3-D audio and tactile cues.

Cue	**Correctly Localized (percent)**		**Error (number of clock positions)**	
	Mean	***SD***	***Mean***	***SD***
3-D Audio	33.5	8.9	1.7	.23
Tactile	82.4	8.5	1.0	.08

Reliance on the visual cue to determine the clock position of targets in the Visual+3-D Audio mode may have reduced the potential for an otherwise large increase in errors in target localization and a decrease in the percentage of hits. As in the present study, the results of an initial

investigation of unimodal cues by Glumm et al. (2005) did not reveal a significant difference between the Visual and the 3-D Audio modes in time to first shot or the percentage of hits. However, targets were easy to detect in the initial study, even at high rates of slew, and the participants may have merely slewed in the direction indicated by the cue rather than attempt to localize the target to a specific clock position. In a following study (i.e., Glumm et al., 2006), targets were well camouflaged and more difficult to detect. A comparison of target acquisition performance between the two investigations revealed a significant interaction of study and cue condition for both time to first shot and the percentage of targets hit. In the 2006 investigation, mean time to first shot in the 3-D Audio mode increased by more than 3 seconds from that of the 2005 study. In the Visual mode, time to first shot increased by only 1 second, and no difference was found between the two studies in time to first shot in the Spatial Language condition. In the 2005 study, 100% of the targets presented were hit in the 3-D Audio mode, but this percentage dropped to 93% in the 2006 investigation. By comparison, 100% of the targets presented in the Visual and Spatial Language modes were hit in both studies.

6.1.5 Comparison of Target Acquisition Performance Between Unimodal and Multimodal Cues

Table 5 shows target acquisition performance in the Visual, Spatial Language, 3-D Audio, and Tactile modes measured across communications modalities in Glumm et al. (2006). Table 6 shows those data obtained in the current investigation in the Visual Only and each of the four multimodal cue conditions. In Glumm et al. (2006), reaction times were slower in the Spatial Language mode than in the other cue conditions, but no significant difference was found between the Spatial Language and Visual modes in time to slew to target or time to first shot. Time to slew and time to first shot were faster in the Spatial Language and Visual modes than in the Tactile and the 3-D Audio conditions. No difference was found between the latter two modes in reaction time, but time to slew to target and time to first shot were faster in the Tactile mode than in the 3-D Audio condition. Although the 3-D Audio cue still proved to be better than no cue at all, time to acquire targets was slower and the percentage of hits achieved was lower than in the Spatial Language, Visual, and Tactile modes.

Table 5. Target acquisition performance in each cue condition across communications modalities (Glumm et al., 2006).

	Unimodal Cues			
Performance Measure	**Visual**	**Spatial Language**	**3-D Audio**	**Tactile**
Reaction Time	0.87	1.16	0.88	0.91
Slew Time	5.82	5.63	8.19	6.80
Time to First Shot	6.68	6.79	9.05	7.71
Percent Hits	99.5	100	92.8	98.3

Table 6. Target acquisition performance in each cue condition across communications modalities (present study).

	Unimodal and Multimodal Cues			
Performance Measure	**Visual Only**	**Performance Measure**	**Visual Only**	**Performance Measure**
Reaction Time	0.84	Reaction Time	0.84	Reaction Time
Slew Time	5.45	Slew Time	5.45	Slew Time
Time to First Shot	6.29	Time to First Shot	6.29	Time to First Shot
Percent Hits	99.2	Percent Hits	99.2	Percent Hits

Table 7 shows the differences in target acquisition performance between Glumm et al. (2006) and the present investigation. As can be seen, there was relatively little change in performance between the Visual and the Visual Only [1] cue conditions, as might be expected. Any improvements in target acquisition performance that might have been influenced by the pairing of the visual cue with the spatial language [2], 3-D audio [3], and tactile [4] cues were not much greater than the differences in performance found between the Visual and the Visual Only condition [1]. Differences in time to slew between the Spatial Language mode and the Visual+Spatial Language [5] and Spatial Language+Tactile conditions [6] were also similar to differences found between the Visual and the Visual Only modes [1]. However, the decrease in reaction time appeared larger and seems to have influenced the decrease in time to first shot. Comparisons of performance in the 3-D Audio mode with performance in the Visual+3-D Audio condition [7] yielded the most dramatic differences in time to slew, time to first shot, and the percentage of hits. Improvements were also found on these measures when performance in the Tactile mode was compared with performance in the Visual+Tactile [8] and the Spatial Language+Tactile conditions [9]; however, these improvements were not as large as those found for the Visual+3-D Audio mode [7].

Table 7. Differences between cue conditions in target acquisition performance in Glumm et al. (2006) and the present study.

Glumm et al. (2006)	**Cue Conditions**								
	Visual				**Spatial Language**		**3-D Audio**	**Tactile**	
Present Study	Visual Only [1]	Visual+ *Spatial Language* [2]	Visual+*3-D Audio* [3]	Visual+ *Tactile* [4]	*Visual+* Spatial Language [5]	Spatial Language+ *Tactile* [6]	*Visual+* 3-D Audio [7]	*Visual+* Tactile [8]	*Spatial Language+* Tactile [9]
Reaction Time	- 0.03	- 0.01	- 0.10	- 0.06	- 0.30	- 0.23	- 0.11	- 0.10	+ 0.02
Slew Time	- 0.37	- 0.60	- 0.40	- 0.26	- 0.41	- 0.30	- 2.77	- 1.24	- 1.47
Time to First Shot	- 0.39	- 0.60	- 0.49	- 0.32	- 0.71	- 0.53	- 2.86	- 1.35	- 1.45
Percent Hits	- 0.30	+ 0.05	- 0.30	- 0.80	0	0	+ 6.40	+ 0.40	+ 1.70

The problems experienced by participants in localizing targets to a specific clock position using the 3-D audio cue are believed to be related to the generic HRTFs used to spatialize the auditory signals. People respond differently to different HRTFs, and the participants' ability to localize targets may have been improved if the HRTFs had been tailored to each participant. Generic HRTFs were also used to spatialize the audio cues in a study by Tannen et al. (2000) of adaptive

and non-adaptive auditory and visual displays of target location in a flight task. The researchers found that the percentage of correct detections was significantly lower in the No Cue and 3-D Audio cue conditions than in conditions where a visual cue or a visual+3-D audio cue was provided. The researchers also attributed the poor performance of the 3-D audio display, in part, to the non-individualized HRTFs.

The difficulty that participants experienced in determining the clock position of the target using the 3-D audio cues appeared to influence their ranking of the Visual+3-D Audio mode as last with respect to how helpful the cue condition was in locating targets. The Visual Only condition was ranked in next-to-last place. Although the visual cue was capable of indicating the direction to slew as well as the clock position of the target, the participants believed that they would need an alternate mode for display of target location information when they are outside their vehicle or in the open hatch position where a visual display is not available. In addition, by itself, a visual cue may at times go unnoticed, as was observed on several occasions during the present study. The Spatial Language+Tactile condition followed the Visual Only condition in the ranking. There were advantages and disadvantages to both cues in this pair. The spatial language cue clearly identified the clock position of the target, but some processing appeared to be needed to translate the words into a direction to slew to acquire the target in the shortest distance. The tactile cue provided more immediate information about slewing direction than the verbal cue, but some participants believed that it took more time to determine the clock position of the target with the tactile cue than it did with the visual or spatial language cue. Little to no difference was found in the ranking of the Visual+Spatial Language and the Visual+Tactile modes. The participants liked the visual cue, and some noted that it was similar to an indicator they currently have on their vehicle that shows the orientation of the main gun with respect to the vehicle chassis. However, many participants claimed that they relied more on the spatial language cue because they were more accustomed to receiving target location information in that format. There were also those who liked the tactile display because it provided a good indication of the direction to slew, but some still questioned the effectiveness of the tactile cue in conveying information in the vibration environment that is typical of the ground combat vehicle.

6.2 Information Recalled from SITREP Communications (Secondary Task)

For this study, it was hypothesized that more information would be recalled from the SITREP communications in the Visual Only condition because there would be fewer sensory signals that would distract from the capture and rehearsal of the information presented. However, contrary to expectations, the analysis did not reveal a significant effect of cue condition on performance of the secondary task.

Although 75% of the participants believed that they were able to recall more information from the SITREP communications when the information was presented auditorily, no differences were found between the auditory or visual communications in the amount of information recalled, and no interaction was found between cue condition and communications modality.

6.3 Subjective Workload

Ratings of frustration tended to be higher in the Spatial Language+Tactile mode than in either the Visual+3-D Audio or the Visual+Tactile conditions. This may reflect the extent to which participants relied on the verbal cue in the former mode and problems they may have encountered in determining the direction to slew. Ratings of frustration tended to decrease, as did reaction times, when the 3-D audio cue was added to the visual cue; however, neither ratings of frustration nor reaction times were significantly affected when the tactile cue was added to the visual cue. These findings may suggest that the 3-D audio cue is more effective than the visual cue or the tactile cue in capturing attention and denoting slewing direction.

Ratings of effort were significantly lower in the Visual+Tactile mode when communications were presented auditorily, whereas no differences were found between communications modalities in the other cue conditions. This may suggest that the auditory communications were less disruptive to the tactile cue than they were to the other cue conditions. Ratings of effort increased significantly in the Visual+Tactile mode when the SITREP communications were presented visually. When communications were visual, ratings of effort in the Visual+Tactile condition were lower than those in the Visual Only mode but not significantly different from ratings of effort in the other cue conditions. Although no significant differences were found in performance between communications modalities in the Visual+Tactile mode, these findings may imply that there is less potential for interference between auditory communications and the tactile cue.

7. Conclusions and Recommendations

The results of this study did not demonstrate a significant improvement in time to first shot or the percentage of hits in multimodal cue conditions beyond that achieved in the Visual Only mode. Reaction times to target location cues were faster when the visual cue was paired with the 3-D audio cue than in all other cue conditions, but no significant differences were found between the Visual+3-D Audio mode and the Visual Only condition for any other measure of target acquisition performance.

Reaction times were slower in modes that included a spatial language cue than in modes where the visual cue was paired with the 3-D audio or the tactile cue. The slower reaction times to the spatial language cues may reflect the time involved in “converting the utterance linguistically into meaning” (Loomis et al., 2002) and into a mental image of the target’s location with respect to the 12 o’clock position. However, differences in reaction times did not result in significant differences in time to first shot or the percentage of hits.

A comparison of performance in a previous assessment of unimodal cues (i.e., Glumm et al., 2006) with the results of the current study also indicated that none of the measures of target acquisition

performance were significantly improved beyond the unimodal Visual condition by the addition of the spatial language, 3-D audio, or tactile cue to the visual cue. However, reaction times were faster than those in the unimodal Spatial Language condition when the visual or the tactile cue was added to the spatial language cue. Even more dramatic were improvements beyond the unimodal 3-D Audio and Tactile conditions that were found in time to slew, time to first shot, and the percentage of hits when the visual cue was added to the 3-D audio or the tactile cue and when the tactile cue was supplemented with the spatial language cue.

The findings of this study suggest that the 3-D audio and tactile cues can provide a more intuitive indication about the direction to slew than the spatial language cue and may be more successful than the visual and the verbal cues in capturing the participants' attention amid visual and verbal communications. However, the results also indicate that the 3-D audio and tactile technologies used in this and the previous assessment may be less effective than the visual and the spatial language cues in denoting the exact location of the target.

Although the percentage of hits tended to be lower in the Visual+Tactile mode than in the Visual+Spatial Language condition, the participants ranked both modes equally as most helpful in locating targets. The Visual Only condition was ranked second only to the Visual+3-D Audio mode as least helpful.

Although the multimodal cues did not significantly improve performance beyond that of the Visual Only condition, additional cues may be essential from an operational perspective. The M1 and M2 commanders and gunners who participated in this study stated that a visual cue about target location should be integrated into their sight picture, as it was in the present investigation. However, they also emphasized that there must be an alternate means for presenting information about target location when they are in the open-hatch position or outside their vehicle where a visual display is not available. Multimodal cues indicating target location may also be necessary to capture attention during periods of high workload or sustained operations when Soldiers are multitasking or fatigued.

When choosing a second cue to pair with a visual cue about target location, one must consider those tasks that might be performed concurrently with the target acquisition task. Although no interactions were found between cue condition and communications modality in the present study, the analysis did indicate that reaction times to target location cues were slower when the communications were presented visually than when they were presented auditorily. This finding emphasizes the importance of avoiding overload of the visual channel by the addition of visual communications to an already visually intensive target acquisition task. However, the potential also exists for auditory cues to be lost among the frequent verbal exchanges, radio transmissions, and perpetual din that are characteristic of a combat vehicle environment. Although the 3-D audio and the tactile cues might seem to be promising candidates, more research and development of these technologies are needed to enhance the fidelity of these cues in denoting the location of targets. The problems experienced by participants in localizing targets using the

3-D audio cue are believed to be attributable, in part, to the generic HRTFs that were used to spatialize the auditory signals. The participants' ability to localize targets may have been improved if the HRTFs had been tailored to each participant. The participants appeared to have less difficulty in localizing targets with the tactile cue than with the 3-D audio cue, but at least one participant perceived that it took less time to determine the clock position of the target with the spatial language cue than with the tactile cue. Thus, in the near term, a more realizable technique for providing information about target location might be to pair a visual cue with a spatial language cue. In addition, the spatial language cue should be presented to one or the other earphone of the crew member's helmet to indicate whether the target lies closer to the left or right of the main gun.

Although the technology needed to detect targets and provide target location cues is not currently defined, research to quantify the advantages that such information can provide and to identify presentation techniques that yield the greater benefit is an important first step. For the purpose of the present study, it was assumed that such a technology would detect 100% of the targets with no false alarms. However, it is recognized that the translation of information from sensors and other intelligence sources into reliable, high fidelity sensory cues about enemy position poses a significant challenge in which failure to ensure the dependability of such cues can present a host of concerns and other research issues. Future studies will examine the effects of the reliability of cues about target location on visual scanning (i.e., eye and head movements), target detection performance, and perceived workload.

8. References

Begault, D. R. Challenges to the Successful Implementation of 3-D Sound. *Journal of the Audio Engineering Society* **1991**, *39* (11), 864-870.

Begault, D. R. Head-up Auditory Displays for Traffic Collision Avoidance System Advisories: A Preliminary Investigation. *Human Factors* **1993**, *35* (4), 707-717.

Boff, K. R.; Lincoln, J. E. *Engineering Data Compendium: Human Perception and Performance: Volume III.* Armstrong Aerospace Medical Research Laboratory: Wright-Patterson AFB, OH, 1988.

Calhoun, G. L.; Draper, M. H.; Ruff, H. A.; Fontejon, J. V. Utility of a Tactile Display in Cueing Faults. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 2144-2148). Santa Monica, CA: Human Factors and Ergonomics Society, 2002.

Calhoun, G. L.; Draper, M. H.; Ruff, H. A.; Fontejon, J. V.; Guilfoos, B. J. Evaluation of Tactile Alerts for Control Station Operation. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 2118-2122). Santa Monica, CA: Human Factors and Ergonomics Society, 2003.

Calhoun, G. L.; Fontejon, J. V.; Draper, M. H.; Ruff, H. A.; Guilfoos, B. J. Tactile Versus Aural Redundant Alert Cues for UAV Control Applications. In *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting* (pp. 137-141). Santa Monica, CA: Human Factors and Ergonomics Society, 2004.

Chiasson, J.; McGrath, B. J.; Rupert, A. H. *Enhanced Situation Awareness in Sea, Air, and Land Environments* (ADP013874). Naval Aerospace Medical Research Lab: Pensacola, FL, 2003.

Colavita, F. B. Human Sensory Dominance. *Perception and Psychophysics* **1974**, *16*, 409-412.

Glumm, M. M.; Kehring, K. L.; White, T. L. *Effects of Visual and Auditory Cues About Threat Location on Target Acquisition and Attention To Auditory Communications;* ARL-TR-3560; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2005.

Glumm, M. M.; Kehring, K. L.; White, T. L. *Effects of Tactile, Visual, and Auditory Cues About Threat Location on Target Acquisition and Attention to Visual and Auditory Communications;* ARL-TR-3863; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2006.

Haas, E. C.; Pillalamarri, R. S.; Stachowiak, C. C.; Lattin, M. A. *Audio Cues to Assist Visual Search in Robotic System Operator Control Unit Displays*; ARL-TR-3632; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2005.

Hart, S. G.; Staveland, L. E. Development of a Multi-Dimensional Workload Rating Scale: Results of Empirical and Theoretical Research. In P.A. Hancock & N. Meshkati (Eds.). *Human Mental Workload.* Amsterdam, The Netherlands: Elsevier, 1988.

Headquarters, Department of the Army. *Use of Volunteers as Subjects of Research*; AR 70-25; Washington, DC, 1990.

Krausman, A. S.; Pettit, R. A.; Elliott, L. R. *Effects of Redundant Alerts on Platoon Leader Performance and Decision Making*; ARL-TR-3999; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2007.

Lockyer, B. *Operation Manual for the MIT Wireless Tactile Control Unit*. Massachusetts Institute of Technology: Cambridge, MA, 2004.

Loomis, J. M.; Lippa, Y.; Klatzky, R. L.; Golledge, R. G. Spatial Updating of Locations Specified by 3-D Sound and Spatial Language. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **2002**, *28,* 335-345.

McKinley, R. L.; Erickson, M. A.; D'Angelo, W. R. Three-dimensional Auditory Displays: Development, Applications, and Performance. *Aviation, Space, and Environmental Medicine* **1994**, *5,* A31–A38.

McKinley, R. L.; D'Angelo, W. R.; Haas, M. W.; Perrot, D. R.; Nelson, W. T.; Hettinger, L. J.; Brickman, B. J. An Initial Study of the Effects of Three-dimensional Auditory Cueing on Visual Target Detection. *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting,* 119-124, Santa Monica, CA, 1995.

Mitchell, D. K.; Samms, C.; Glumm, M.; Krausman, A.; Brelsford, M.; Garrett, L. *Improved Performance Research Integration Tool (IMPRINT) Model Analyses in Support of the Situational Understanding as an Enabler for Unit of Action Maneuver Team Soldiers Science and Technology Objective (STO) in Support of Future Combat Systems (FCS)*; ARL-TR-3405; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2004.

Simpson, B. D.; Brungart, D. S.; Gilkey, R. H.; Cowgill, J. L.; Dallman, R. C.; Green, R. F.; Youngblood, K. L.; Moore, T. J. 3-D Audio Cueing for Target Identification in a Simulated Flight Task. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting,* 1836-1840, Santa Monica, CA, 2004.

Tannen, R. S.; Nelson, W. T.; Bolia, R. S.; Haas, M. W.; Warm, J. S., Hettinger, L. J. Adaptive Integration of Head-Coupled Multi-Sensory Displays for Target Localization. *Proceedings of the IEA 2000/HFES Congress* **2000**, *3,* 77-80.

U.S Army. *Hearing Conservation;* U.S. Army Pamphlet 40-501; Department of Defense: Washington, DC, 1991.

Van Erp, J. B.; Meppelink, R.; Van Veen, H. A. *A Touchy Car* (TNO-TM-02-B001). Human Factors Research Institute: TNO Soesterberg, The Netherlands, 2002.

Van Erp, J. B.; Veltman, J. A.; Van Veen, H. A.; Oving, A. B. *Tactile Torso Display as Countermeasure To Reduce Night Vision Goggles Induced Drift* (ADP013888). Human Factors Research Institute: TNO Soesterberg, The Netherlands, 2003.

Wickens, C. D. Multiple Resources and Performance Prediction. *Theoretical Issues in Ergonomics Science* **2002**, *3* (2), 159-177.

Wickens, C. D.; Hollands, J. G. Attention, Time Sharing, and Workload. *Engineering Psychology and Human Performance.* Prentice Hall: New Jersey, 2000.

INTENTIONALLY LEFT BLANK

Appendix A. Situation Report (SITREP[7]) Example

1. Your objective is 500 meters east of your current assembly point.
2. Your platoon is to arrive at objective Brown at 0500 hours.
3. Threats on the way to your objective include an ambush
4. Closest supporting unit is infantry, 100 meters from objective.
5. Call sign of closest supporting unit is Charlie 1.
6. Friendly artillery unit is 2 kilometers north of airport.
7. Call sign for air support unit is Eagle 1.
8. Drop off point for squad is Peach Hill, 100 meters south of objective.
9. Closest enemy unit to the objective is armor.
10. Closest enemy unit is 200 meters from your objective.
11. Closest enemy unit is company-size.
12. Closest enemy unit to your objective is resupplying.
13. Enemy infantry in countryside are armed with RPGs.
14. Enemy has placed landmines near your objective
15. Enemy has placed obstacles near objective along Church Road.

[7]The 20 items of information underlined are those that changed in deriving the 25 SITREPs. Only 10 of these items were changed in each version of the SITREP.

INTENTIONALLY LEFT BLANK

Appendix B. SITREP Questionnaire Example

Participant # : ______ Cue/Communications Mode: __/__ Target Set : ___

SITREP Questionnaire

Please answer the following questions based on the SITREP you heard during this last target set. Each answer is one or two words. The number in parentheses after the question indicates the number of words in the answer. Examples of one-word answers are "armor" or "NBC". Examples of two-word answers are "mechanized infantry" or "Charlie Company". An answer that requires two words can also include a number. Examples of two-word answers with a number are "200 meters" or "Charlie 35". Each answer is worth 2 points. If you do not provide an answer to a question, or the answer is wrong, you will lose 2 points. If you omit a word from an answer that requires two words, or if one of the words in your answer is wrong, you will lose 1 point.

Question	Answer *(Please PRINT)*
1. What is the name of your objective? (1)	________________________
2. How far is the objective from your current assembly point? (2)	________________________
3. At what time is your platoon to arrive at the objective? (2)	________________________
4. What type of threat might you encounter on the way to your objective? (1)	________________________
5. What is the name of the place where the squad is to dismount? (2)	________________________
6. What is the call sign of your closest supporting unit? (2)	________________________
7. What type of enemy unit is closest to the objective? (1)	________________________
8. How far is the closest enemy unit from the objective? (2)	________________________
9. In what activity is the enemy unit that is closest to your objective currently engaged? (1)	________________________
10. What is the size of the enemy unit that is closest to your objective? (1)	________________________

INTENTIONALLY LEFT BLANK

Appendix C. NASA-TLX

Rating Scale Definitions

Title	End Points	Descriptions
MENTAL DEMAND	Low/High	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	Low/High	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	Low/High	How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
PERFORMANCE	Perfect/Failure	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
EFFORT	Low/High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION LEVEL	Low/High	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Effort or **Performance**	**Temporal Demand** or **Frustration**
Temporal Demand or **Effort**	**Physical Demand** or **Frustration**
Performance or **Frustration**	**Physical Demand** or **Temporal Demand**
Physical Demand or **Performance**	**Temporal Demand** or **Mental Demand**

Frustration or **Effort**	**Performance** or **Mental Demand**
Performance or **Temporal Demand**	**Mental Demand** or **Effort**
Mental Demand or **Physical Demand**	**Effort** or **Physical Demand**
Frustration or **Mental Demand**	

RATING SCALE SHEET

MENTAL DEMAND

Low **High**

PHYSICAL DEMAND

Low **High**

TEMPORAL DEMAND

Low **High**

PERFORMANCE

Good **Poor**

EFFORT

Low **High**

FRUSTRATION

Low **High**

Appendix D. Demographic Questionnaire

Please answer the following questions. The information you provide will be considered privileged and will be protected.

1. Participant No: ____

2. Age: _________ years

3. Rank: __________________

4. Military Occupational Specialty (MOS): ________________

5. Time in Service: ______years ________ months

6. Time in MOS: _______years ________ months

7. Are you left- or right-handed?

   Left-Handed [ ]   Right-Handed [ ]

8. How often do you play video- or computer games?

Never	[ ]
Sometimes	[ ]
Often	[ ]
All the time	[ ]

9. Vehicle Experience and Crew Positions Held *(Check all that apply):*

	Commander	Gunner	Loader	Driver
Bradley Fighting Vehicle	[ ]	[ ]	[NA]	[ ]
M1 tank	[ ]	[ ]	[ ]	[ ]
Other *(Please specify)* __________	[ ]	[ ]	[ ]	[ ]

10. Combat Experience:

Geographic Area *(Check all that apply)*		Duration of Tour	Did you see combat? *(Circle either Yes or No)*		**If "Yes"** *(Please specify duty position during combat)*
Bosnia	[ ]	___years ___months	Yes	No	_____________
Afghanistan	[ ]	___years ___months	Yes	No	_____________
Iraq 1 (Desert Storm)	[ ]	___years ___months	Yes	No	_____________
Iraq 2 (Iraqi Freedom)	[ ]	___years ___months	Yes	No	_____________

INTENTIONALLY LEFT BLANK

Appendix E. Post-Test Questionnaire

Participant No.: ____

1. Please rank *each* of the five cue conditions from 1 to 5 based on the order that you thought each was helpful in finding targets (1 = Most Helpful, 5 = Least Helpful):

Baseline	[ ]
Visual	[ ]
Verbal	[ ]
3-D Audio	[ ]
Tactile	[ ]

2. For each of the five cue conditions, in what communications modality was it easier to remember the information contained in the SITREPs?

Baseline	Auditory [ ]	Visual [ ]	No Difference [ ]
Visual	Auditory [ ]	Visual [ ]	No Difference [ ]
Verbal	Auditory [ ]	Visual [ ]	No Difference [ ]
3-D Audio	Auditory [ ]	Visual [ ]	No Difference [ ]
Tactile	Auditory [ ]	Visual [ ]	No Difference [ ]

3. Did the *auditory* SITREPs interfere with your ability to acquire targets in any of the five cue conditions?

Baseline	Yes [ ]	No [ ]
Visual	Yes [ ]	No [ ]
Verbal	Yes [ ]	No [ ]
3-D Audio	Yes [ ]	No [ ]
Tactile	Yes [ ]	No [ ]

4. Did the *visual* SITREPs interfere with your ability to acquire targets in any of the five cue conditions?

Baseline	Yes [ ]	No [ ]
Visual	Yes [ ]	No [ ]
Verbal	Yes [ ]	No [ ]
3-D Audio	Yes [ ]	No [ ]
Tactile	Yes [ ]	No [ ]

5. How often did you use the azimuth indicator in each of the five cue conditions?

Baseline	Never [ ]	Sometimes [ ]	Often [ ]	All the time [ ]
Visual	Never [ ]	Sometimes [ ]	Often [ ]	All the time [ ]
Verbal	Never [ ]	Sometimes [ ]	Often [ ]	All the time [ ]
3-D Audio	Never [ ]	Sometimes [ ]	Often [ ]	All the time [ ]
Tactile	Never [ ]	Sometimes [ ]	Often [ ]	All the time [ ]

NO. OF COPIES	ORGANIZATION
1 (PDF ONLY)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA 8725 JOHN J KINGMAN RD STE 0944 FORT BELVOIR VA 22060-6218
1	US ARMY RSRCH DEV & ENGRG CMD SYSTEMS OF SYSTEMS INTEGRATION AMSRD SS T 6000 6TH ST STE 100 FORT BELVOIR VA 22060-5608
1	DIRECTOR US ARMY RESEARCH LAB IMNE ALC IMS 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	DIRECTOR US ARMY RESEARCH LAB AMSRD ARL CI OK TL 2800 POWDER MILL RD ADELPHI MD 20783-1197
2	DIRECTOR US ARMY RESEARCH LAB AMSRD ARL CS OK T 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR M DR M STRUB 6359 WALKER LANE SUITE 100 ALEXANDRIA VA 22310
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR ML J MARTIN MYER CENTER RM 2D311 FT MONMOUTH NJ 07703-5601
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MZ A DAVISON 199 E 4TH ST STE C TECH PARK BLDG 2 FT LEONARD WOOD MO 65473-1949
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MD T COOK BLDG 5400 RM C242 REDSTONE ARSENAL AL 35898-7290

NO. OF COPIES	ORGANIZATION
1	COMMANDANT USAADASCH ATTN AMSRD ARL HR ME A MARES 5800 CARTER RD FT BLISS TX 79916-3802
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MO J MINNINGER BLDG 5400 RM C242 REDSTONE ARSENAL AL 35898-7290
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MM DR V RICE-BERG BLDG 4011 RM 217 1750 GREELEY RD FT SAM HOUSTON TX 78234-5094
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MG R SPINE BUILDING 333 PICATINNY ARSENAL NJ 07806-5000
1	ARL HRED ARMC FLD ELMT ATTN AMSRD ARL HR MH C BURNS BLDG 1467B ROOM 336 THIRD AVENUE FT KNOX KY 40121
1	ARMY RSCH LABORATORY - HRED AVNC FIELD ELEMENT ATTN AMSRD ARL HR MJ D DURBIN BLDG 4506 (DCD) RM 107 FT RUCKER AL 36362-5000
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MK MR J REINHART 10125 KINGMAN RD FT BELVOIR VA 22060-5828
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MV HQ USAOTC S MIDDLEBROOKS 91012 STATION AVE ROOM 111 FT HOOD TX 76544-5073
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MY M BARNES 2520 HEALY AVE STE 1172 BLDG 51005 FT HUACHUCA AZ 85613-7069

NO. OF COPIES	ORGANIZATION
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MP D UNGVARSKY BATTLE CMD BATTLE LAB 415 SHERMAN AVE UNIT 3 FT LEAVENWORTH KS 66027-2326
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MJK J HANSBERGER JFCOM JOINT EXPERIMENTATION J9 JOINT FUTURES LAB 115 LAKEVIEW PARKWAY SUITE B SUFFOLK VA 23435
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MQ M R FLETCHER US ARMY SBCCOM NATICK SOLDIER CTR AMSRD NSC SS E BLDG 3 RM 341 NATICK MA 01760-5020
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MY DR J CHEN 12423 RESEARCH PARKWAY ORLANDO FL 32826
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MS MR C MANASCO SIGNAL TOWERS 118 MORAN HALL FORT GORDON GA 30905-5233
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MU M SINGAPORE 6501 E 11 MILE RD MAIL STOP 284 BLDG 200A 2ND FL RM 2104 WARREN MI 48397-5000
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MF MR C HERNANDEZ BLDG 3040 RM 220 FORT SILL OK 73503-5600
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MW E REDDEN BLDG 4 ROOM 332 FT BENNING GA 31905-5400
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MN R SPENCER DCSFDI HF HQ USASOC BLDG E2929 FORT BRAGG NC 28310-5000

NO. OF COPIES	ORGANIZATION
1	ARMY G1 ATTN DAPE MR B KNAPP 300 ARMY PENTAGON ROOM 2C489 WASHINGTON DC 20310-0300
1	US ARMY NATICK SOLDIER CTR FUTURE FORCE WARRIOR ATTN AMSRB NSC W C BLACKWELL NATICK MA 01760-5020
1	UNIV OF CENTRAL FLORIDA DEPT OF PSYCHOLOGY R GILSON 4000 CENTRAL FLORIDA BLVD ORLANDO FL 32816-1390
	<u>ABERDEEN PROVING GROUND</u>
1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL CI OK (TECH LIB) BLDG 4600
1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL CI OK TP S FOPPIANO BLDG 459
2	US ARMY ABERDEEN TEST CTR ATTN DEPY DIR FUTURE FORCE DIR CHIEF TRANSF DIV P TENNANT BLDG 400
1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL HR MR F PARAGALLO BLDG 459
3	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL HR SB M GLUMM KATHY KEHRING TIM WHITE BLDG 459

NO. OF COPIES	ORGANIZATION
1	JAN BF VAN ERP TNO HUMAN FACTORS DEPT OF HUMAN INTERFACE PO BOX 23 3769 ZG SOESTERBERG THE NETHERLANDS